JOURNAL OF THE ROYAL SOCIETY Interface

J. R. Soc. Interface (2007) **4**, 949–961
doi:10.1098/rsif.2007.1031
Published online 8 May 2007

REVIEW

The pluses and minuses of $\mathcal{R}_0$

M. G. Roberts*

Centre for Mathematical Biology, Institute of Information and Mathematical Sciences, Massey University, Private Bag 102 904, North Shore Mail Centre, Auckland, New Zealand

The concept of the basic reproduction number ($\mathcal{R}_0$) occupies a central place in epidemic theory. The value of $\mathcal{R}_0$ determines the proportion of the population that becomes infected over the course of a (modelled) epidemic. In many models, (i) an endemic infection can persist only if $\mathcal{R}_0>1$, (ii) the value of $\mathcal{R}_0$ provides a direct measure of the control effort required to eliminate the infection, and (iii) pathogens evolve to maximize their value of $\mathcal{R}_0$. These three statements are not universally true. In this paper, some exceptions to them are discussed, based on the extensions of the SIR model.

Keywords: mathematical epidemiology; infectious diseases; basic reproduction number; type reproduction number; backward bifurcation

1. INTRODUCTION

The basic reproduction number ($\mathcal{R}_0$) of an infection is defined by Diekmann & Heesterbeek (2000) as the 'expected number of secondary cases per primary case in a virgin population'. In this context, a virgin population is one that is fully susceptible to the infection in question. If $\mathcal{R}_0>1$, then an epidemic is expected to occur following the introduction of infection and if $\mathcal{R}_0<1$, then the number infected in the population is expected to decrease following introduction and the infection will be eliminated over time. These concepts are well known, see for example the books by Anderson & May (1991) and Diekmann & Heesterbeek (2000). For a formal mathematical discussion of the concept see Diekmann *et al.* (1990), for a history see Heesterbeek (2002) and for a recent review of the formulation, estimation and use of $\mathcal{R}_0$ in deterministic epidemic models see Heffernan *et al.* (2005). Although Kermack & McKendrick (1927) did not use the term 'basic reproduction number', or any of its synonyms, they discussed the concept of an epidemic threshold and derived the final size equation (2.2). In their later papers, Kermack & McKendrick discussed the problem of endemicity. Their combined contribution has been reviewed by Diekmann *et al.* (1995), who also remind readers that the SIR model is a special case, not *the Kermack–McKendrick model*.

For an endemic infection, the value $\mathcal{R}_0=1$ defines a threshold. Below this threshold, an infectious agent will not invade and become established in a previously uninfected population. Above this threshold, the pathogen can invade. It is usually the case that if $\mathcal{R}_0>1$, then an endemic infection will persist in the host population, and if $\mathcal{R}_0$ is subsequently reduced below 1 by control measures, then the pathogen will be eliminated. The basic reproduction number is not the only threshold parameter that may be used to derive qualitative information about the dynamics of endemic infections. For some examples, the critical community size has been used (Anderson & May 1991). For other infections, notably those with obligatory multi-host life cycles, the definition of $\mathcal{R}_0$ in the model has not been applied consistently (Roberts & Heesterbeek 2003). For these cases, it is always useful to formulate the definition of $\mathcal{R}_0$ in words and then translate this into system parameters. For models expressed as dynamical systems, the threshold between stability and instability of the infection-free steady state may be determined by linearizing the system about the steady state and finding a combination of parameters that makes all the eigenvalues of the Jacobian matrix negative (or if complex numbers, have negative real parts; see texts on dynamical systems, for example, Strogatz 1994; Jordan & Smith 1999). This process does not, in itself, define which combination of parameters should be called $\mathcal{R}_0$ or has the correct biological definition.

In §2, we review the SIR model. We present the SIR epidemic model and the final size equation, which determines the proportion of the population that gets infected in an epidemic as a function of $\mathcal{R}_0$. We then present the SIR endemic model and show that the infection can persist only if $\mathcal{R}_0>1$, and the value of $\mathcal{R}_0$ determines the effort required to eliminate an infection from the population. We introduce the SIR endemic model with two pathogens and reciprocal immunity, and show that the pathogens cannot coexist and the

*m.g.roberts@massey.ac.nz

One contribution of 20 to a Theme Issue 'Cross-scale influences on epidemiological dynamics: from genes to ecosystems'.

Received 2 March 2007
Accepted 16 April 2007

pathogen that excludes the other is that with the greater value of $\mathcal{R}_0$. This implies that a pathogen evolving through mutation will maximize its basic reproduction number. All of these results may be found in the literature. In §3, we review some extensions of the SIR endemic model in which an infection may persist even if $\mathcal{R}_0<1$. These are (i) the SIR model where those recovered (in the R-class) are susceptible to infection, and perhaps more susceptible than those in the S-class, (ii) the SEI model with exogeneous infection, (iii) an SIR model with nonlinear transmission, and (iv) a model with a carriage state. We then discuss structured epidemic models and the type reproduction number ($\mathcal{T}$). We present an example where $\mathcal{R}_0$ by itself does not determine the control effort required to eliminate infection, but $\mathcal{T}$ does. Finally, we discuss a model for a fatal infectious disease in a population with a variable host density. We show that for this model, two variants of a pathogen can coexist and that evolution of the pathogen is not necessarily in favour of the maximization of $\mathcal{R}_0$.

2. THE SIR MODEL

The so-called SIR model is a special case of the Kermack–McKendrick model (Kermack & McKendrick 1927; Diekmann & Heesterbeek 2000). A population of constant size is divided into three compartments, with the proportion susceptible or infectious at time t equal to $s(t)$ or $i(t)$, respectively. The proportion 'removed' from the infection process is $r(t)=1-s(t)-i(t)$.

2.1. The SIR epidemic model

In many situations, the time-scale of the infection is much faster than the demographic time-scale. For example, epidemics of measles or influenza take place over a matter of months, with each infected individual participating for less than a month (e.g. Roberts & Tobias 2000; Roberts *et al.* 2007). In a human population, the demographic time-scale is measured in tens of years. Hence, as an approximation, we can ignore population dynamics and consider the SIR epidemic model

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t}&=-\beta si,\\ \frac{\mathrm{d}i}{\mathrm{d}t}&=\beta si-\gamma i.\end{aligned}\tag{2.1}$$

The parameters β and γ are rates and have units time^{-1}. If the population is fully susceptible, then $s=1$, and those infected infect others at an expected rate β for an expected time $1/\gamma$. Hence $\mathcal{R}_0=\beta/\gamma$. Solutions to equations (2.1) may be plotted in an (s, i) phase plane by observing that

$$i(t)+s(t)-\frac{1}{\mathcal{R}_0}\log s(t)=\text{const.}$$

An epidemic occurs if $\mathcal{R}_0 s(0)>1$, and in the (s, i) plane this may be represented as a curve from the point $(s(0), 0)$ to the point $(s(\infty), 0)$. (Strictly, this is a heteroclinic trajectory. For details, see texts on dynamical systems, for example, Strogatz 1994; Jordan & Smith 1999.) The proportion infected in an epidemic is $z=s(0)-s(\infty)$, where

$$\mathcal{R}_0+\frac{1}{z}\log\left(1-\frac{z}{s(0)}\right)=0.\tag{2.2}$$

It is easy to show that z is an increasing function of $\mathcal{R}_0$ (Diekmann & Heesterbeek 2000). Equation (2.2) is often known as the *final size equation* and provides an approximation to the number infected in an epidemic as determined by more complicated models (Roberts *et al.* 2007). For example, if the whole population is initially susceptible ($s(0)=1$) and $\mathcal{R}_0=2$, then $z=0.7968$, and nearly 80% of the population is infected in the course of an epidemic.

2.2. The SIR endemic model

If the time-scale of the infection is of the same order as the demographic time-scale, we need to consider the SIR endemic model. One example would be tuberculosis, where the infection may persist for years (Blower *et al.* 1995; Gomes *et al.* 2004). In its simplest form, with constant and equal birth and death rates (μ), the equations are

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t}&=\mu-\beta si-\mu s,\\ \frac{\mathrm{d}i}{\mathrm{d}t}&=\beta si-\gamma i-\mu i.\end{aligned}\tag{2.3}$$

For this model, those infected infect others for an expected time $1/(\gamma+\mu)$, hence $\mathcal{R}_0=\beta/(\gamma+\mu)$. Equations (2.3) have two steady states, $(s, i)=(1, 0)$ and $(s, i)=(s^*, i^*)=(1/\mathcal{R}_0, \mu(1-s^*)/(\gamma+\mu))$. As s^* is a proportion, the endemic steady state (s^*, i^*) makes biological sense only if $\mathcal{R}_0>1$, and it is then globally attracting (appendix A.1). If $\mathcal{R}_0<1$, then the infection-free steady state $(1, 0)$ is globally attracting; hence, the objective of an infection eradication programme must be to reduce $\mathcal{R}_0$ below 1. If this is to be achieved by vaccinating a proportion v of the population at birth, then $(2.3)_1$ becomes

$$\frac{\mathrm{d}s}{\mathrm{d}t}=(1-v)\mu-\beta si-\mu s,$$

and no endemic steady state exists if $v>1-1/\mathcal{R}_0$. If eradication is to be achieved by treating infecteds at a rate θ, then $(2.3)_2$ becomes

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\beta si-\gamma i-\mu i-\theta i,$$

and no endemic steady state exists if $\theta/(\gamma+\mu)>\mathcal{R}_0-1$. Hence, the value of $\mathcal{R}_0$ provides a direct measure of the control effort required to achieve eradication.

2.3. The SIR endemic model with competition

Now consider the case where two pathogens exist and there is total reciprocal immunity: infection or having been infected with one excludes infection with the other. Such a situation could arise, for example, with two competing strains of the same pathogen.

The population is now divided into four compartments, with $s+i_1+i_2+r=1$, and

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t} &= \mu-\beta_1 s i_1-\beta_2 s i_2-\mu s,\\ \frac{\mathrm{d}i_1}{\mathrm{d}t} &= \beta_1 s i_1-\gamma_1 i_1-\mu i_1,\\ \frac{\mathrm{d}i_2}{\mathrm{d}t} &= \beta_2 s i_2-\gamma_2 i_2-\mu i_2.\end{aligned} \tag{2.4}$$

For equations (2.4), coexistence, in the form of steady states with both i_1^* and i_2^* non-zero, is not possible (unless $\mathcal{R}_0^1=\mathcal{R}_0^2$, see appendix A.2). The steady state with pathogen 1 present and pathogen 2 absent ($i_1^*\neq 0, i_2^*=0$) exists when $\mathcal{R}_0^1>1$ and is stable if $\mathcal{R}_0^1>\mathcal{R}_0^2$. Hence, if a pathogen undergoes a mutation that increases $\mathcal{R}_0$, then the new variant will outcompete the original variant.

Recall that for the SIR endemic model (2.3), $\mathcal{R}_0=\beta/(\mu+\gamma)$. We can write $\beta=p\kappa$, where p is the probability of transmission given contact and κ is the contact rate. It is reasonable that $\kappa'(p)\leq 0$, if somebody is perceived to be infectious they and others may take steps to reduce contact. Hence, we could write either $\beta'(p)\leq 0$ or $\beta'(p)\geq 0$. In addition, either of $\gamma'(p)\leq 0$ or $\gamma'(p)\geq 0$ could be reasonable, a highly infectious pathogen may persist in the host for longer or may result in a short severe illness. Dieckmann *et al.* (2002) use p as a measure of *matching virulence* and γ as a measure of *aggressive virulence*. With either measure, maximizing $\mathcal{R}_0$ is not the same as maximizing virulence.

In this section, we have demonstrated that for an SIR model, an epidemic will occur if $\mathcal{R}_0>1$ and the number infected in an epidemic is an increasing function of $\mathcal{R}_0$. We have also shown that (i) an endemic infection can persist only if $\mathcal{R}_0>1$, (ii) the value of $\mathcal{R}_0$ provides a direct measure of the control effort needed to eliminate the infection, and (iii) a pathogen evolves to maximize $\mathcal{R}_0$. In the following sections, we discuss various extensions to the SIR model that result in exceptions to these three conclusions.

3. SUBCRITICAL PERSISTENCE

In this section, we review some models in which a pathogen may invade an uninfected host population only if $\mathcal{R}_0>1$, but an endemic infection may persist for some values of $\mathcal{R}_0<1$. The first two examples are special cases of a model with exogenous reinfection and treatment effects proposed for tuberculosis by Feng *et al.* (2000). In their model, the population had a constant recruitment rate into the susceptible (S) class, proceeded to the exposed (E) class upon exposure to infection and then to the infectious (I) class. Those in the infectious class were treated at a constant rate and joined a recovered class, labelled T for treatment. As described, this is a standard SEIR model. However, there were two added features. If an individual in the E-class was reexposed to infection, then their transition rate to the I-class increased; hence, reexposure accelerated the development of infectiousness. If an individual in the T-class was reexposed to infection, then they could return to the E-class; hence, the treated class remained susceptible to infection. Each of the first two models described below is a special case of this system with just one of these features, and in each, it is possible that an infection persists even though $\mathcal{R}_0<1$. Hence, the model of Feng *et al.* (2000) has two features that can lead to a backward bifurcation and subcritical persistence of infection.

These two examples involve, in some sense, a force of infection that is not proportional to the proportion of the population infectious as it is in the SIR model. In the third example, this is made explicit in that instead of $\lambda=\beta i$ we take $\lambda=\beta i(1+h(i))$ for some non-decreasing function h. This option is used as an approximation for situations where the force of infection is somehow 'dose dependent', and a higher prevalence of infection is correlated with a higher exposure to the pathogen on infection. Gomes *et al.* (2005) suggest that this mechanism may apply to the epidemiology of measles, polio and foot-and-mouth disease among others, and that infectiousness with tuberculosis may increase with the intensity and frequency of exposure.

The fourth example of a model that may lead to the subcritical persistence is based on one analysed by Medley *et al.* (2001) and used to describe the dynamics of hepatitis B. This SEIR model had the features that a proportion of births to mothers in a carrier class are also carriers and a proportion of the infectious class become carriers upon recovery. The carriers continue to be infectious and the proportion of infectives who become carriers is a non-decreasing function of the force of infection. We simplify the model by ignoring the exposed class and vaccination, which was also a feature, and assuming vertical transmission to be negligible. We show that the maintenance of carriers in the host population may lead to persistence after $\mathcal{R}_0$ is reduced below 1. In §4.1, we consider the same model but with vertical transmission included as an example of a model with structure.

For the examples discussed in this section, we use $\mathcal{R}_0$ as the primary bifurcation parameter. This enables us to distinguish different qualitative behaviour of the models for different parameter values. In some instances, this leads us to what could strictly be regarded as loose terminology, for example suggesting that $\mathcal{R}_0$ increases from one value to another where we intend to contrast situations where $\mathcal{R}_0$ may be lower or higher, respectively. Hence, in figures 1 and 2, where $\mathcal{R}_0$ is allowed to vary along the horizontal axes, this is done by changing β and keeping all other parameters constant. In §§3.1, 3.2, 3.4 and 4.1, we use $\mathcal{P}$ as a secondary bifurcation parameter. At each use, $\mathcal{P}$ is defined and the definition should be regarded as local to §3.1, §3.2, or §§3.4 and 4.1 as appropriate.

3.1. A model with a susceptible R-class

This is a special case of the models examined by Feng *et al.* (2000), which also included an exposed class (see §3.2), and Gomes *et al.* (2005), which also included nonlinear transmission (see §3.3). It is equivalent to that discussed by Safan *et al.* (2006). The recovered

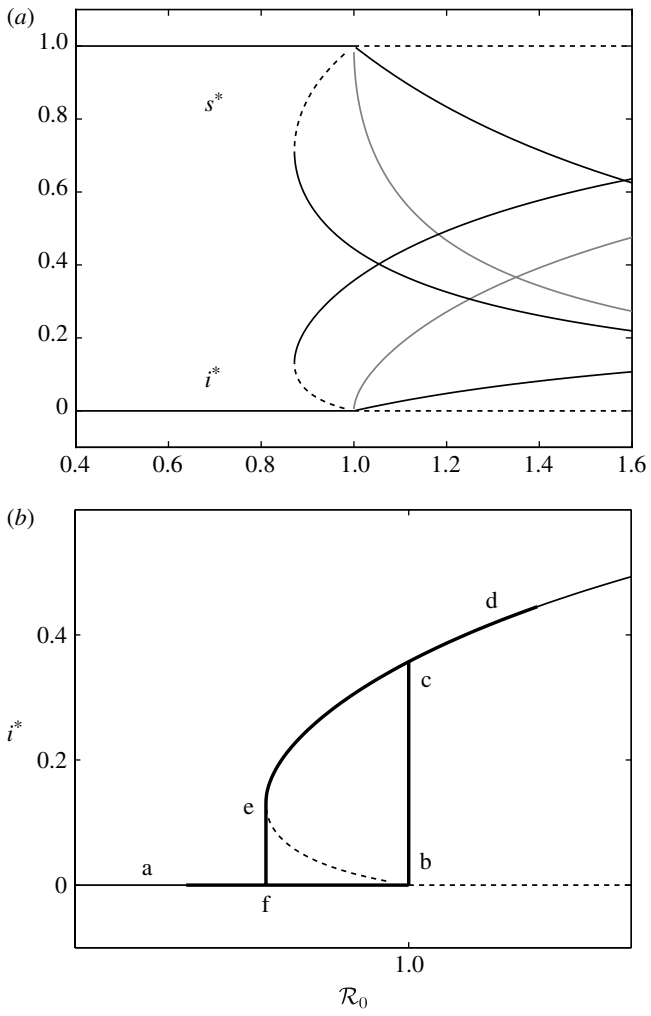

Figure 1. (*a*) Bifurcation diagram for the model with a susceptible *R*-class. Curves are top s^* and bottom i^* as functions of $\mathcal{R}_0$. Broken lines signify unstable steady states and unbroken lines stable steady states. Curves are plotted for $\mathcal{P}=0$ (forward bifurcation), $\mathcal{P}=\mathcal{P}_{\rm crit}$ (shown in grey) and $\mathcal{P}=2\mathcal{P}_{\rm crit}$ (backward bifurcation), together with the trivial steady state $s=1$, $i=0$. Other parameters are $\mu=0.02$ and $\gamma=0.05$. (*b*) An enlargement of the $(\mathcal{R}_0, i^*)$ curve for $\mathcal{P}=2\mathcal{P}_{\rm crit}$, showing the possibility of a hysteresis effect. For explanation of the lettering see the text.

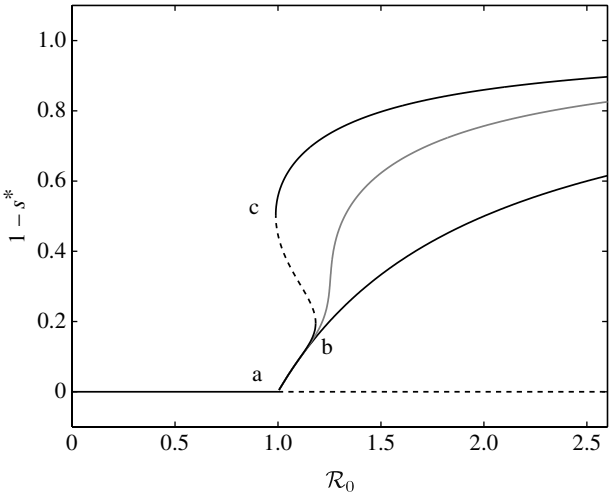

Figure 2. Bifurcation diagram for the carriage model. The curves are $1-s^*$ as a function of $\mathcal{R}_0$, broken lines signify unstable steady states and unbroken lines stable steady states. Curves are plotted for $\mathcal{P}=0$ (lower curve), $\mathcal{P}=\mathcal{P}_{\rm crit}$ (shown in grey) and $\mathcal{P}=2\mathcal{P}_{\rm crit}$ (upper curve), together with the trivial steady state $s=1$. Parameter values are $\gamma=0.05$, $\mu=1/70$, $\delta=0.025$ and $k=1$. The function $q(x)=\exp(-x^{-1})$, where $x=\beta(i^*+\mathcal{P}c^*)/\mu$ (appendix A.6) and $\mathcal{P}_{\rm crit}=0.8$. For explanation of the lettering see the text.

class is assumed to be susceptible to infection and, possibly, more susceptible than the susceptible class (hence, in this example it is not removed from the epidemic). We write

$$\begin{aligned}\frac{{\rm d}s}{{\rm d}t} &= \mu-\beta si-\mu s,\\ \frac{{\rm d}i}{{\rm d}t} &= \beta si+\mathcal{P}\beta ri-(\mu+\gamma)i,\end{aligned} \tag{3.1}$$

with $r(t)=1-s(t)-i(t)$ and $\mathcal{R}_0=\beta/(\mu+\gamma)$ as before. The parameter $\mathcal{P}$ is the ratio of the susceptibility of the recovered class to that of the susceptible class. If $\mathcal{R}_0>1$, then the endemic steady state (s^*, i^*) is the solution of a quadratic equation, leading to the possibility of multiple steady states and bistability (appendix A.3). An endemic steady state exists and is locally stable for some values of $\mathcal{R}_0$ less than 1 if $\mathcal{P}>\mathcal{P}_{\rm crit}=1+\mu/\gamma$. A bifurcation diagram for this model is shown in figure 1*a*.

Figure 1*a* demonstrates how the steady-state values (s^*, i^*) of equations (3.1) depend on $\mathcal{R}_0$ and $\mathcal{P}$. For all values of $\mathcal{P}$, the infection-free steady state is stable for $\mathcal{R}_0<1$ and unstable for $\mathcal{R}_0>1$. When $\mathcal{P}<\mathcal{P}_{\rm crit}$ and $\mathcal{R}_0>1$, there is a unique globally stable endemic steady state (s^*, i^*). However, if $\mathcal{P}>\mathcal{P}_{\rm crit}$, the locus of steady states leaves the bifurcation point at $\mathcal{R}_0=1$ to the left (backwards) with an unstable steady state (shown as broken), then changes direction and becomes stable. At $\mathcal{P}=\mathcal{P}_{\rm crit}$, these curves leave the bifurcation point vertically (shown in grey) and this separates two qualitatively different behaviours. For $\mathcal{P}<\mathcal{P}_{\rm crit}$, the stable steady-state solutions of equations (3.1) change continuously as $\mathcal{R}_0$ is increased or decreased through the critical value $\mathcal{R}_0=1$. The behaviour for $\mathcal{P}>\mathcal{P}_{\rm crit}$ is further illustrated in figure 1*b*. As $\mathcal{R}_0$ is increased through $\mathcal{R}_0=1$, the stable steady-state value of i increases discontinuously, moving from the infection-free solution to the endemic solution i^*. In figure 1*b*, this is shown by the line starting with $\mathcal{R}_0<1$ at 'a', through the critical value $\mathcal{R}_0=1$ where it jumps from the infection-free steady state at 'b' to the endemic steady state at 'c', then further increasing $\mathcal{R}_0$ increases the value of i^* to, say, 'd'. If $\mathcal{R}_0$ were then decreased, the value of i^* would decrease along the solid curve shown in figure 1*b* from 'd' to 'e', remaining positive at and below $\mathcal{R}_0=1$ in the portion from 'c' to 'e'. At 'e', another discontinuous change would occur as the only steady state possible is $i=0$ and the system moves to 'f'. Hence, increasing and then decreasing $\mathcal{R}_0$ through the bifurcation point could cause a hysteresis effect with two discontinuous changes, for example moving around figure 1*b* in the sequence 'f,b,c,e,f'.

3.2. A model with exogenous infection

For the second special case of the model by Feng *et al.* (2000), we reinstate the exposed class, ignore the treated (recovered) class and restore the mechanism of exogenous reinfection: exposure to infection accelerates transition from E to I at a rate proportional to the force of infection, that is from ν to $\nu+\mathcal{P}\beta i$ for some constant $\mathcal{P}$. Feng *et al.* (2000) suggest that $\mathcal{P}<1$ might be appropriate for tuberculosis, but $\mathcal{P}>1$ could be appropriate for HIV. The equations are written as

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t}&=\mu-\beta si-\mu s,\\ \frac{\mathrm{d}e}{\mathrm{d}t}&=\beta si-\mathcal{P}\beta ei-(\mu+\nu)e,\\ \frac{\mathrm{d}i}{\mathrm{d}t}&=\mathcal{P}\beta ei+\nu e-\mu i,\end{aligned}\tag{3.2}$$

with $s(t)+e(t)+i(i)=1$. Hence, this model is two dimensional and, if $\mathcal{P}=0$, is a standard SEI model. For all values of $\mathcal{P}$, the basic reproduction number is $\mathcal{R}_0=\beta\nu/(\mu(\mu+\nu))$. Equations (3.2) have a bifurcation diagram similar in appearance to figure 1*a*, with a backward bifurcation at $(\mathcal{R}_0, i)=(1, 0)$ occurring if $\mathcal{P}>\mathcal{P}_{\mathrm{crit}}=\nu(\mu+\nu)/\mu^2$ (appendix A.4).

3.3. A model with nonlinear transmission

Gomes *et al.* (2005) provide an extensive analysis of epidemic models with nonlinear transmission and partial or temporary immunity. Here, we focus on the nonlinear or enhanced transmission. The equations are written as

$$\begin{aligned}\frac{\mathrm{d}s}{\mathrm{d}t}&=\mu-\lambda s-\mu s,\\ \frac{\mathrm{d}i}{\mathrm{d}t}&=\lambda s-(\mu+\gamma)i,\end{aligned}\tag{3.3}$$

with $r(t)=1-s(t)+i(t)$ as before, but with $\lambda=\beta i(1+h(i))$. We assume that $h(0)=0$ and $h'(i)\geq 0$, hence we still have $\mathcal{R}_0=\beta/(\mu+\gamma)$. Equations (3.3) also have a bifurcation diagram similar in appearance to figure 1*a*, with a backward bifurcation at $(\mathcal{R}_0, i^*)=(1, 0)$ if $h'(0)>1+\gamma/\mu$ (appendix A.5).

3.4. A model with a carrier class

For our simple example, a proportion q of the infectious class enters a carrier class upon recovery and carriers may still transmit infection. The others (proportion $1-q$) enter the removed class. The model may be written as

$$\begin{aligned}\frac{\mathrm{d}i}{\mathrm{d}t}&=\lambda s-(\mu+\gamma)i,\\ \frac{\mathrm{d}c}{\mathrm{d}t}&=q\gamma i-(\mu+\delta)c,\\ \frac{\mathrm{d}r}{\mathrm{d}t}&=(1-q)\gamma i+\delta c-\mu r,\end{aligned}\tag{3.4}$$

with $s(t)=1-i(t)-c(t)-r(t)$. The force of infection is $\lambda=\beta(i+\mathcal{P}c)$ for some constant $\mathcal{P}$. Hence, $\mathcal{P}$ is the ratio of the infectivity of an individual in the carrier class to the infectivity of an individual in the infectious class. Motivated by hepatitis B, where the proportion of those who become carriers upon recovery is higher when the force of infection is higher (Medley *et al.* 2001), we take q to be a non-decreasing function of λ (appendix A.6). If $q(0)=0$, then $\mathcal{R}_0=\beta/(\mu+\gamma)$ as in the previous models discussed. If not, then even at low infection prevalences, a proportion $q\gamma/(\mu+\gamma)$ of infectives become carriers and infect susceptibles at the rate $\mathcal{P}\beta$ for a mean duration of $1/(\mu+\delta)$ years. Adding this contribution, $\mathcal{R}_0=(\beta/(\mu+\gamma))(1+\mathcal{P}\gamma q(0)/(\mu+\delta))$. The bifurcation diagram is similar in appearance to figure 1*a*, with a backward bifurcation at $(\mathcal{R}_0, i)=(1, 0)$ occurring if $\mathcal{P}>\mathcal{P}_{\mathrm{crit}}=(\mu+\delta)/((q'(0)-q(0))\gamma)$ (appendix A.6).

Hence, we have shown that there are at least four mechanisms that would allow an endemic infection to persist even though $\mathcal{R}_0<1$: (i) enhanced susceptibility of those recovered, (ii) exogenous reinfection of those exposed, (iii) nonlinear transmission of infection, and (iv) the existence of a carrier class.

4. STRUCTURED MODELS

We now consider the situations where more than one type of infective can be identified. By *type* we mean *type at birth*, an attribute that makes two infectives epidemiologically distinct (Diekmann & Heesterbeek 2000). For example, in the previous model specified by equations (3.4), although infectives and carriers transmit infection at different rates, all carriers were previously infectives. Hence, they are not distinct types. However, if a proportion of those born to carriers are themselves carriers, we then have two types: those infected through contact and those infected at birth. This is the next example considered.

In general, suppose we can identify n distinct types of infective. We construct the $n\times n$ next-generation matrix $\boldsymbol{K}$ by defining the element $K_{j\ell}$ to be the expected number of infectives of type j that would arise from an infective of type ℓ in a fully susceptible population. Hence, $\boldsymbol{K}$ is a matrix whose entries are reproduction numbers of some sort and $\mathcal{R}_0$ is defined as the spectral radius (largest eigenvalue) of $\boldsymbol{K}$. For a full discussion see Diekmann & Heesterbeek (2000).

4.1. Type at infection

Consider the model for an infectious disease with carriage described by equations (3.4) and add a second transmission route by allowing that a proportion k of those whose mothers are carriers are born carriers. The system becomes

$$\begin{aligned}\frac{\mathrm{d}i}{\mathrm{d}t}&=\lambda s-(\mu+\gamma)i,\\ \frac{\mathrm{d}c}{\mathrm{d}t}&=k\mu c+q\gamma i-(\mu+\delta)c,\\ \frac{\mathrm{d}r}{\mathrm{d}t}&=(1-q)\gamma i+\delta c-\mu r,\end{aligned}\tag{4.1}$$

with $s(t)=1-i(t)-c(t)-r(t)$, $\lambda=\beta(i+\mathcal{P}c)$ and q an increasing function of λ as before. We can now identify two types of infectious hosts. Define an infectious host of

Type 1 by having been infected through contact and an infectious host of Type 2 by having been infected at birth. The next-generation matrix is calculated as follows.

$K_{1,1}$: infectious hosts infect others at rate β and remain infectious for an expected $1/(\gamma+\mu)$ years. A proportion $q\gamma/(\gamma+\mu)$ of infectious hosts recover to become carriers, who then infect others at rate $\mathcal{P}\beta$, and remain carriers for an expected $1/(\delta+\mu)$ years.

$K_{1,2}$: the mean number of susceptible hosts infected by each carrier is $\mathcal{P}\beta/(\delta+\mu)$.

$K_{2,1}$: on recovery, a proportion $q\gamma/(\gamma+\mu)$ of infectious hosts become carriers and remain carriers for $1/(\delta+\mu)$ years. All hosts give birth at rate μ, and a proportion k of those born to carriers are carriers.

$K_{2,2}$: the mean number of carriers born to each carrier is $k\mu/(\delta+\mu)$.

The basic reproduction number $\mathcal{R}_0$ is the spectral radius of

$$\boldsymbol{K}=\begin{pmatrix}\dfrac{\beta}{\gamma+\mu}+\dfrac{q(0)\gamma}{\gamma+\mu}\dfrac{\mathcal{P}\beta}{\delta+\mu} & \dfrac{\mathcal{P}\beta}{\delta+\mu}\\ \dfrac{q(0)\gamma}{\gamma+\mu}\dfrac{k\mu}{\delta+\mu} & \dfrac{k\mu}{\delta+\mu}\end{pmatrix}. \tag{4.2}$$

An example of the behaviour of this model is presented in figure 2. The figure shows three bifurcation curves, plotting steady-state values of $1-s^*$ against $\mathcal{R}_0$. This is equal to the proportion of the population who have been infected ($i^*+c^*+r^*$). For all values of $\mathcal{P}$ and $\mathcal{R}_0$, we have the trivial steady state $s=1$, which is stable for $\mathcal{R}_0<1$ and unstable for $\mathcal{R}_0>1$. For $\mathcal{P}=0$, the carriers do not transmit infection and $\mathcal{R}_0=\max(\beta/(\gamma+\mu), k\mu/(\delta+\mu))$. The second term is always less than 1, hence when above threshold $\mathcal{R}_0=\beta/(\gamma+\mu)$ (appendix A.6). The diagram shows a unique non-trivial steady state for $\mathcal{R}_0>1$. At a critical value of $\mathcal{P}$, there is an inflection in the bifurcation curve (shown in grey) and this separates two qualitatively different behaviours. If $\mathcal{P}>\mathcal{P}_{\mathrm{crit}}$, then there is the possibility of multiple steady states for a given value of $\mathcal{R}_0$. This does not correspond to a backward bifurcation at $\mathcal{R}_0=1$, as the curve leaves the trivial solution in a *forward* direction at 'a'. However, at some $\mathcal{R}_0>1$, the curve moves through the vertical at 'b', goes *backward* and the steady state becomes unstable. This continues until 'c', where the curve reverses direction again and the steady state becomes stable. For the example shown in figure 2, $\mathcal{R}_0<1$ at 'c', hence persistence of infection is possible for some $\mathcal{R}_0<1$. For values of $\mathcal{R}_0$ between that at 'c' and that at 'b', there are two stable steady states, and the limiting state for large t is determined by initial conditions. This model could still exhibit hysteresis: beginning from $\mathcal{R}_0<1$ and $s=1$ then increasing $\mathcal{R}_0$, the endemic steady state is realized at $\mathcal{R}_0=1$ 'a' and $1-s^*$ increases until 'b', where there is a discontinuous transition to the upper (stable) steady state. If $\mathcal{R}_0$ is then decreased, the steady-state solution maintains the upper curve until 'c', where for lower values of $\mathcal{R}_0$ only the trivial steady state exists.

4.2. A model with two hosts

Infections that have obligatory two-host life cycles include those transmitted by vectors, such as dengue and malaria, and similar models may be used for those transmitted sexually. Motivated by the simplest example of a sexually transmitted infection in a population of constant size, we analyse a two-host SIS model

$$\begin{aligned}\frac{\mathrm{d}i_{\mathrm{f}}}{\mathrm{d}t}&=\beta_{\mathrm{m}}s_{\mathrm{f}}i_{\mathrm{m}}-\gamma_{\mathrm{f}}i_{\mathrm{f}},\\ \frac{\mathrm{d}i_{\mathrm{m}}}{\mathrm{d}t}&=\beta_{\mathrm{f}}s_{\mathrm{m}}i_{\mathrm{f}}-\gamma_{\mathrm{m}}i_{\mathrm{m}},\end{aligned} \tag{4.3}$$

with $s_{\mathrm{f}}=1-i_{\mathrm{f}}$ and $s_{\mathrm{m}}=1-i_{\mathrm{m}}$. A stable endemic steady state exists for $\mathcal{R}_0>1$, where (appendix A.7)

$$\mathcal{R}_0=\sqrt{\frac{\beta_{\mathrm{f}}\beta_{\mathrm{m}}}{\gamma_{\mathrm{f}}\gamma_{\mathrm{m}}}}.$$

The ratio $\beta_{\mathrm{f}}/\gamma_{\mathrm{f}}$ is the expected number of males that would be infected by an infectious female if all males were susceptible, with a similar interpretation for $\beta_{\mathrm{m}}/\gamma_{\mathrm{m}}$. Hence, $\mathcal{R}_0$ is the geometric mean of two $\mathcal{R}_0$-like quantities. For models of this type, the definition of the basic reproduction number has not been used consistently in the literature. For example, compare Feng & Velasco-Hernandez (1997) and Soewono & Supriatna (2001) with Esteva & Vargas (1998, 2000).

Now consider the model described by equations (4.3), but extended to include two competing variants of the pathogen. The equations become

$$\begin{aligned}\frac{\mathrm{d}i_{\mathrm{f}j}}{\mathrm{d}t}&=\beta_{\mathrm{m}j}s_{\mathrm{f}}i_{\mathrm{m}j}-\gamma_{\mathrm{f}j}i_{\mathrm{f}j},\\ \frac{\mathrm{d}i_{\mathrm{m}j}}{\mathrm{d}t}&=\beta_{\mathrm{f}j}s_{\mathrm{m}}i_{\mathrm{f}j}-\gamma_{\mathrm{m}j}i_{\mathrm{m}j},\end{aligned} \tag{4.4}$$

for $j=1, 2$, with $s_{\mathrm{f}}+i_{\mathrm{f}1}+i_{\mathrm{f}2}=s_{\mathrm{m}}+i_{\mathrm{m}1}+i_{\mathrm{m}2}=1$. For a steady state with both variants present $\mathcal{R}_0^1=\mathcal{R}_0^2$, and the state with variant 1 only present is stable if $\mathcal{R}_0^1>\mathcal{R}_0^2$ (appendix A.8). Hence, this model suggests that as a sexually transmitted infection mutates, the fittest variant is the one that maximizes $\mathcal{R}_0$.

The results presented so far for the two-host SIS model suggest a similarity with the results for the endemic SIR model. The direct connection between $\mathcal{R}_0$ and the control effort required to eliminate infection has been lost however, and, in particular, when we introduce a third group of hosts. We now extend the model, letting the subscript p denote a high-risk female group. We assume that those in this group are $\mathcal{P}$ times more likely to transmit infection to males than those in the other female group, where $\mathcal{P}$ is a constant. Equations (4.3) become

$$\begin{aligned}\frac{\mathrm{d}i_{\mathrm{p}}}{\mathrm{d}t}&=\beta_{\mathrm{p}}s_{\mathrm{p}}i_{\mathrm{m}}-\gamma_{\mathrm{p}}i_{\mathrm{p}},\\ \frac{\mathrm{d}i_{\mathrm{f}}}{\mathrm{d}t}&=\beta_{\mathrm{m}}s_{\mathrm{f}}i_{\mathrm{m}}-\gamma_{\mathrm{f}}i_{\mathrm{f}},\\ \frac{\mathrm{d}i_{\mathrm{m}}}{\mathrm{d}t}&=\beta_{\mathrm{f}}s_{\mathrm{m}}(\mathcal{P}i_{\mathrm{p}}+i_{\mathrm{f}})-\gamma_{\mathrm{m}}i_{\mathrm{m}}.\end{aligned} \tag{4.5}$$

We define $\mathcal{R}_0$ as the spectral radius of the next-generation matrix

$$\boldsymbol{K} = \begin{pmatrix} 0 & 0 & \dfrac{\beta_\mathrm{p}}{\gamma_\mathrm{m}} \\ 0 & 0 & \dfrac{\beta_\mathrm{m}}{\gamma_\mathrm{m}} \\ \mathcal{P}\dfrac{\beta_\mathrm{f}}{\gamma_\mathrm{p}} & \dfrac{\beta_\mathrm{f}}{\gamma_\mathrm{f}} & 0 \end{pmatrix}.$$

Hence

$$\mathcal{R}_0 = \sqrt{K_{13}K_{31} + K_{23}K_{32}}.$$

Suppose now that a proportion v of the high-risk group is protected from infection, by vaccination or some other prophylactic measure. Then, the basic reproduction number under vaccination becomes

$$\mathcal{R}_\mathrm{v} = \sqrt{(1-v)K_{13}K_{31} + K_{23}K_{32}}. \tag{4.6}$$

Alternatively, if the high-risk group is treated at a rate θ, the equation for that group becomes

$$\frac{\mathrm{d}i_\mathrm{p}}{\mathrm{d}t} = \beta_\mathrm{p}s_\mathrm{p}i_\mathrm{m} - \gamma_\mathrm{p}i_\mathrm{p} - \theta i_\mathrm{p}, \tag{4.7}$$

and the basic reproduction number under treatment becomes

$$\mathcal{R}_\mathrm{c} = \sqrt{\frac{\gamma_\mathrm{p}}{\gamma_\mathrm{p}+\theta}K_{13}K_{31} + K_{23}K_{32}}. \tag{4.8}$$

Neither $\mathcal{R}_\mathrm{v}$ nor $\mathcal{R}_\mathrm{c}$ on their own specify the control effort necessary to eliminate infection from the population.

4.3. The type reproduction number

When calculating $\mathcal{R}_0$ in a multi-host situation, one is effectively averaging over all host types by taking the eigenvalue of $\boldsymbol{K}$. When applying control measures unevenly across the host types, the previous direct relationship between $\mathcal{R}_0$ and control effort is lost. This difficulty can often be overcome by using the type reproduction number ($\mathcal{T}$), defined as the number of secondary infections of Type 1 that would arise from a single primary infection of Type 1 introduced to a fully susceptible population (Roberts & Heesterbeek 2003; Heesterbeek & Roberts 2007). In our example (equations (4.5)), we define $\mathcal{T}$ for the high-risk group (Type 1, subscript p) by

$$\mathcal{T} = [\boldsymbol{K}(\boldsymbol{I} - (\boldsymbol{I}-\boldsymbol{P})\boldsymbol{K})^{-1}]_{11}, \tag{4.9}$$

with $P_{11}=1$, $P_{ij}=0$ otherwise. Hence $\mathcal{T} = K_{13}K_{31}/(1-K_{23}K_{32})$. It has been shown (Roberts & Heesterbeek 2003) that $\mathcal{T}<1$ if and only if $\mathcal{R}_0<1$. Moreover, if a proportion v of the high-risk group is protected from infection, then the type reproduction number becomes $\mathcal{T}_\mathrm{v}=(1-v)\mathcal{T}$, and to eliminate the infection we require $v>1-1/\mathcal{T}$. Moreover, if the infecteds in the high-risk group are treated at a rate θ (equation (4.7)), then the type reproduction number becomes $\mathcal{T}_\mathrm{c}=(\gamma_p/(\gamma_p+\theta))\mathcal{T}$, and to eliminate the infection we require $\theta/\gamma_\mathrm{p} > \mathcal{T}-1$. Given $\mathcal{T}$, the expressions for $\mathcal{T}_\mathrm{v}$ and $\mathcal{T}_\mathrm{c}$ provide direct measures of the critical control effort. These are much easier to apply than those derived from the basic reproduction number under control (equations (4.6) and (4.8)). This methodology is readily extended to the situation where control is applied to more than one host type (Roberts & Heesterbeek 2003).

In summary, for the multi-host model, if $\mathcal{R}_0<1$, the pathogen cannot invade the host population and the pathogen evolves to maximize $\mathcal{R}_0$. The value of $\mathcal{R}_0$ is determined by averaging over all host types, but this definition is not applied consistently in the literature and does not directly determine the control effort required to eliminate infection. The type reproduction number ($\mathcal{T}$) focuses on a particular host type and provides a direct measure of the control effort required to eliminate the infection.

5. WILDLIFE INFECTIONS

Let N denote the size of the host population or more properly the number of hosts in a given geographical region (measured as host density). We relax the assumption that the host population size is constant. As a consequence, the host birth and death rates, and the contact rate between hosts, may be functions of $N(t)$. The birth and death rates are non-increasing and non-decreasing functions ($\eta'(N)\leq 0$ and $\mu'(N)\geq 0$), respectively, with a unique solution of the equation $\eta(N)=\mu(N)$ at $N=K$ being the host population carrying capacity.

Recall that we can write the transmission rate (β) as the product of the contact rate (κ) and the probability of transmission given contact (p). The contact rate is now a non-decreasing function of host population density, hence we have $\beta'(N)\geq 0$. Common functional forms are $\beta(N)$ constant, often called frequency-dependent transmission, and $\beta(N)\propto N$, often called density-dependent transmission. Other terms such as mass action, pseudo mass action and standard incidence have appeared in the literature.

A model for a fatal infection in a wild animal population was analysed by Roberts (1996). The equations for the host population density and the proportion of hosts infectious are written as

$$\begin{aligned} \frac{\mathrm{d}N}{\mathrm{d}t} &= (\eta(N) - \mu(N) - \alpha i)N, \\ \frac{\mathrm{d}i}{\mathrm{d}t} &= \beta(N)si - (\alpha s + \eta(N))i, \end{aligned} \tag{5.1}$$

with $s+i=1$. A brief derivation is given in appendix A.9. The basic reproduction number $\mathcal{R}_0$ is calculated at carrying capacity, $\mathcal{R}_0=\mathcal{R}(K)$, where $\mathcal{R}(N)=\beta(N)/(\alpha+\mu(N))$. This model has another threshold quantity: if $\beta(0)/(\alpha+\eta(0))>1$, then a steady state exists with $N=0$ but $i\neq 0$. It may be thought strange that the contact rate could be positive when the host population density is zero, but this may be interpreted as the infection driving the host population to extinction. A unique endemic steady-state solution with $N^*\neq 0$ and $0<i^*<1$ exists if $\mathcal{R}_0>1$ (appendix A.9).

Control procedures may be incorporated in this model as before. If infecteds are treated at a rate θ, no endemic steady state exists if $\theta/(\alpha+\mu(K))>\mathcal{R}_0-1$. If a proportion v of the host population is maintained vaccine immune, then no endemic steady state exists if $v>1-1/\mathcal{R}_0$. If the host species is regarded as a pest, then it is an option to eliminate infection by culling animals, say at a rate χ, and no endemic steady state exists if $\beta(N)/(\alpha+\eta(N))<1$, where $\eta(N)=\mu(N)+\chi$.

We now consider small mutations in the pathogen. With two variants, equations (5.1) become

$$\begin{aligned}\frac{\mathrm{d}N}{\mathrm{d}t} &= (\eta(N)-\mu(N)-\alpha_1 i_1-\alpha_2 i_2)N,\\ \frac{\mathrm{d}i_j}{\mathrm{d}t} &= \beta_j(N)si_j-(\alpha_j+\eta(N)-\alpha_1 i_1-\alpha_2 i_2)i_j,\end{aligned} \tag{5.2}$$

for $j=1, 2$, with $s+i_1+i_2=1$. A similar model (with η constant and $\beta\propto N$) was analysed by Andreasen & Pugliese (1995). Define $\mathcal{R}_j(N)=\beta_j(N)/(\alpha_j+\mu(N))$ for $j=1, 2$. Then, $\mathcal{R}_j(K)$ is the basic reproduction number for pathogen j in the absence of the other pathogen. Clearly, the infection-free steady state is stable if $\mathcal{R}_1(K)<1$ and $\mathcal{R}_2(K)<1$. There is a steady state with pathogen 1 only present if $\mathcal{R}_1(K)>1$, at which $(N, i_1, i_2)=(N_1^*, i_1^*, 0)$. This is stable if $\mathcal{R}_2(N_1^*)<1$ (appendix A.10). Similarly, the steady state with pathogen 2 only present exists if $\mathcal{R}_2(K)>1$ and is stable if $\mathcal{R}_1(N_2^*)<1$. However, there is now a third possibility: a steady state with both pathogens present. Numerical results show that this coexistence steady state exists and is stable whenever both single-pathogen steady states exist but are unstable. This is only the case for a limited range of parameter values. If pathogen variant 1 only is present, then variant 2 can invade if $\mathcal{R}_2(N_1^*)>1$, but this is not the same as $\mathcal{R}_2(K)>\mathcal{R}_1(K)$. In contrast to the model with constant population size, the established parasite has a role in determining a *modified carrying capacity*, N_1^*, and it is not necessary that the pathogen with the largest basic reproduction number excludes the other. The order in which the pathogens are established in the host population matters.

6. DISCUSSION

The basic reproduction number ($\mathcal{R}_0$) remains the single-most useful quantity to calculate or estimate when modelling the population dynamics of an infectious disease. For simple models, and these are sufficient in many practical situations, it is true that (i) the value of $\mathcal{R}_0$ determines whether a pathogen may invade or persist in a host population, (ii) pathogens evolve to maximize $\mathcal{R}_0$, and (iii) the value of $\mathcal{R}_0$ provides a direct measure of the control effort required to eliminate the infection. The purpose of this review has been to draw attention to some situations where these properties do not always apply, and where possible to indicate additional or alternative quantities that define the qualitative epidemiology of the pathogen.

In all of the examples of models for single pathogens that we have discussed, it is true that if $\mathcal{R}_0$ for the pathogen is less than 1, then the pathogen will not invade and infect a previously uninfected and completely susceptible host population. It is often also true that if $\mathcal{R}_0$ for the pathogen is reduced below 1, then a pathogen that was previously present in the host population will be eliminated. The examples based on models proposed for tuberculosis and other infections by Feng *et al.* (2000) and Gomes *et al.* (2005) exhibit backward bifurcations leading to subcritical persistence for some parameter values (figure 1), and it is not sufficient to reduce $\mathcal{R}_0$ below 1 to eliminate the infection. Some of these and other examples have been discussed by van den Driessche & Watmough (2002). The model proposed by Medley *et al.* (2001) could lead to subcritical persistence, but without a backward bifurcation at $\mathcal{R}_0=1$ (figure 2). The paper by Safan *et al.* (2006) is of particular interest here, as it presents a method for determining the control effort required to eliminate an infection from a host population when subcritical persistence may occur.

The definition of $\mathcal{R}_0$ in two-host populations has often led to confusion, even if the threshold property itself has remained the same. For example, in a model for malaria or dengue, we could define a next-generation matrix $\boldsymbol{K}$ with K_{12}, the expected number of humans infected by an infectious mosquito, and K_{21}, the expected number of mosquitoes infected by an infectious human, both in a fully susceptible population. The diagonals of $\boldsymbol{K}$ equal zero, hence $\mathcal{R}_0=\sqrt{K_{12}K_{21}}$. Taking a geometric mean number of humans and mosquitoes seems a strange thing to do, although it is a valid threshold quantity. The type reproduction number $\mathcal{T}=\mathcal{R}_0^2$ is both the expected number of secondary infections in humans that would arise from a primary infected human and the expected number of secondary infections in mosquitoes that would arise from a primary infected mosquito, both in a fully susceptible population. It is $\mathcal{T}$ that provides a direct link with the control effort required to eliminate infection (Heesterbeek & Roberts 2007). A similar problem arises in the definition of a threshold quantity for macroparasites. A naive definition based on individual stages in an obligatory life cycle can result in the value of the analogue of $\mathcal{R}_0$ depending on the level of detail in the model, which is at odds with a property having a biological definition (Heesterbeek & Roberts 1995*a*). Once again, it is the definition that is altered but not the threshold property.

Many models of the evolution of an infectious disease have led to the conclusion that the fittest variant of a pathogen is the one that maximizes $\mathcal{R}_0$. However, it has previously been demonstrated that in situations where a pathogen may change the population dynamics of the host, the order in which the variants are established is important (Dieckmann *et al.* 2002). This was also shown for communities of parasitic helminths in wild animal populations by Roberts & Dobson (1995). Dieckmann & Metz (2006) present examples of simple models of infectious disease where the evolutionarily stable strategy does not maximize $\mathcal{R}_0$. Another

instance where a pathogen may, in theory at least, evolve towards a lower value of $\mathcal{R}_0$ has been discussed by Kao (2006), where the seemingly anomalous behaviour is linked to the ability to exploit heterogeneities in the host population.

In this exposition, we have assumed the environment to be static, and therefore all models discussed have been autonomous. This is often not appropriate. For example, epidemics of measles occur during the school year (Roberts & Tobias 2000) and a model requires a seasonal variation in transmission parameter $\beta(t)$. Heesterbeek & Roberts (1995*b*) addressed this problem by way of Floquet theory, which requires defining a discrete map from year to year. A similar approach was taken for a model with birth pulses by Roberts & Kao (1998). In contrast, Grassly & Fraser (2006) defined a quantity $\bar{\mathcal{R}}_0 = \gamma^{-1}\int_0^1 \beta(t)\mathrm{d}t$, appealing to the result that if $i(t)$ solves the equations $i' = \beta(t)i - \gamma i$ and $\bar{\mathcal{R}}_0 < 1$, then $\lim_{t\to\infty} i(t) = 0$. This does not make $\bar{\mathcal{R}}_0$ equal to the basic reproduction number. The definition of $\mathcal{R}_0$ in a periodic environment is difficult—both of the methods mentioned above assume that the primary case is equally likely to arise at any time during the year. If the infection is seasonal on a global scale, then the introduction of a primary case will be more likely in the high season for the epidemic. Grassly & Fraser (2006) acknowledge this. They also criticize the Floquet approach as requiring that the Poincaré map be calculated and hence the equations must be solved numerically over one time period. Many packages are available that will do this as easily as calculating an eigenvalue, so that is no impediment to the method.

Throughout the exposition, we have also assumed that the population is large and the unit is the individual host. In some applications, it is more appropriate to take the unit as the household (e.g. Ball & Becker 2006) or farm (e.g. Haydon *et al.* 2003) or some other entity. There is increasing interest in modelling epidemics on networks (e.g. Kiss *et al.* 2006; May 2006). In a network model, a node can represent any entity that may be considered susceptible, infectious, etc. and a link between nodes indicates a potential for infectious contact between the two entities. Spatial heterogeneity can be explicitly represented in this way. A useful result for a random network is that if all links represent the same probability of transmission, and $g(j)$ is the number of nodes that have j links, then $\mathcal{R}_0 \propto \mathrm{mean}(g) + \mathrm{var}(g)/\mathrm{mean}(g)$. Homogeneous mixing is equivalent to assuming that all nodes are connected and $\mathrm{var}(g) = 0$, hence heterogeneity in a network serves to increase $\mathcal{R}_0$. A network, by its very nature, is finite in size, so it is not immediately apparent how concepts such as $\mathcal{R}_0$ that are based on large (infinite) populations translate, and how their properties may be adapted to finite structures. In addition, all of the results that we have cited have been for deterministic models. For finite populations, stochastic models are appropriate, but results for these are much more difficult to obtain and they are beyond the scope of this review.

In conclusion, we have reviewed the concept of the basic reproduction number $\mathcal{R}_0$ in simple deterministic models of the dynamics of infectious diseases. We have shown that if $\mathcal{R}_0 < 1$, a pathogen cannot invade a host population, but mechanisms exist by which a pathogen may persist if $\mathcal{R}_0 < 1$. We have also shown that for structured populations, the type reproduction number ($\mathcal{T}$) provides a direct measure of the control effort required to eliminate an infection and has the same threshold property as $\mathcal{R}_0$. And finally, we have shown that pathogens usually evolve to maximize $\mathcal{R}_0$, but this may be modified by host population dynamics.

The author has had many fruitful and stimulating discussions on these topics with Odo Diekmann, Klaus Dietz, Hans Heesterbeek and Hans Metz. Yue Zhao assisted with computations. Comments from four anonymous referees have led to considerable improvements in the paper.

APPENDIX A

A.1. The SIR endemic model

The results quoted require the global asymptotic stability of the steady states. Clearly, if $s = 0$, $\mathrm{d}s/\mathrm{d}t > 0$; if $i = 0$, $\mathrm{d}i/\mathrm{d}t = 0$, and all higher derivatives of i are zero; and if $r = 0$, $\mathrm{d}r/\mathrm{d}t \geq 0$. Hence, if $(s(t), i(0))' \in \Omega$, where the prime signifies transpose, and $\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 : x_1, x_2 \geq 0 \text{ and } x_1 + x_2 \leq 1\}$, then $(s(t), i(t))' \in \Omega$ for $t > 0$. Then, to establish global stability of steady states, with or without treatment and vaccination, Dulac's criterion may be used (Edelstein-Keshet 2005, pp. 327–330). As

$$\frac{\partial}{\partial s}\left(\frac{(1-v)\mu - \beta si - \mu s}{si}\right) + \frac{\partial}{\partial i}\left(\frac{\beta si - \gamma i - \mu i - \theta i}{si}\right) = -\frac{(1-v)\mu}{s^2 i} < 0,$$

no limit cycles exist, and a steady state that is locally stable is also globally stable. The Jacobian matrix corresponding to equations (2.3), with both treatment and vaccination included, is

$$\boldsymbol{J} = \begin{pmatrix} -\beta i - \mu & -\beta s \\ \beta i & \beta s - \gamma - \mu - \theta \end{pmatrix}.$$

The infection-free steady state is $(s, i) = (1-v, 0)$, which is stable if the matrix $\boldsymbol{J}$ evaluated at the steady state has negative eigenvalues, i.e. if $(1-v)\beta < \gamma + \mu + \theta$, which is equivalent to $(1-v)\mathcal{R}_0 < 1 + \theta/(\gamma + \mu)$. The expressions for the critical vaccination and treatment efforts follow. Note that the eigenvalues of $\boldsymbol{J}$ do not define an expression for $\mathcal{R}_0$.

A.2. The SIR endemic model with competition

For equations (2.4), coexistence, in the form of a steady state $(s, i_1, i_2) = (s^*, i_1^*, i_2^*)$ with both i_1^* and i_2^* non-zero, would require $s^* = (\gamma_1 + \mu)/\beta_1 = 1/\mathcal{R}_0^1$ from $(2.4)_2$, $s^* = (\gamma_2 + \mu)/\beta_2 = 1/\mathcal{R}_0^2$ from $(2.4)_3$, and hence $\mathcal{R}_0^1 = \mathcal{R}_0^2$. Consider the steady state with $i_1^* \neq 0$ and $i_2^* = 0$, which exists when $\mathcal{R}_0^1 > 1$. Solutions to (2.4) near to the steady state may be written $(s(t), i_1(t), i_2(t))' = (s^*, i_1^*, 0)' + \boldsymbol{x}(t) + \mathcal{O}(|\boldsymbol{x}|^2)$,

where $\dot{\boldsymbol{x}} = \boldsymbol{J}\boldsymbol{x}$ and

$$\boldsymbol{J} = \begin{pmatrix} -\mu\mathcal{R}_0^1 & -\beta_1 s^* & -\beta_2 s^* \\ \mu(\mathcal{R}_0^1 - 1) & 0 & 0 \\ 0 & 0 & \beta_2\left(\dfrac{1}{\mathcal{R}_0^1} - \dfrac{1}{\mathcal{R}_0^2}\right) \end{pmatrix}.$$

The overdot signifies time derivative. The matrix $\boldsymbol{J}$ is the Jacobian matrix of the system (2.4) evaluated at the steady state $(s(t), i_1(t), i_2(t)) = (s^*, i_1^*, 0)$. If $\boldsymbol{J}$ has eigenvalues, all of which are negative or complex numbers with negative real parts, then the steady state is locally stable. Solutions that begin near to the steady state tend towards it. Otherwise it is unstable. In this example, the eigenvalues of $\boldsymbol{J}$ are equal to the eigenvalues of the leading 2×2 submatrix, which is always stable for $\mathcal{R}_0^1 > 1$, and to the third diagonal entry. Hence, the steady state is locally stable if $\mathcal{R}_0^1 > \mathcal{R}_0^2$ and unstable otherwise.

A.3. A model with a susceptible R-class

Using Dulac's criterion,

$$\frac{\partial}{\partial s}\left(\frac{\mu - \beta si - \mu s}{si}\right) + \frac{\partial}{\partial i}\left(\frac{\beta si + \mathcal{P}\beta ri - (\mu+\gamma)i}{si}\right) = -\frac{\mu}{s^2 i} - \frac{\mathcal{P}\beta}{s} < 0.$$

Hence, no limit cycles exist and the non-trivial steady state is globally stable for $\mathcal{R}_0 > 1$. The steady-state values of s and i solve

$$\mathcal{P}\mathcal{R}_0^2 i^{*2} + \left(1 + \mathcal{P}\frac{\mu}{\mu+\gamma} - \mathcal{P}\mathcal{R}_0\right)\mathcal{R}_0 i^* + \frac{\mu}{\mu+\gamma}(1-\mathcal{R}_0) = 0$$

$$\mathcal{R}_0(\mathcal{P}-1)s^{*2} + \left(1 - \mathcal{P}\frac{\mu}{\mu+\gamma} - \mathcal{P}\mathcal{R}_0\right)s^* + \mathcal{P}\frac{\mu}{\mu+\gamma} = 0.$$

Treating $\mathcal{R}_0$ as a function of s^*, differentiating the last equation and setting $(\mathcal{R}_0, s^*) = (1, 1)$, we obtain

$$\left.\frac{\mathrm{d}\mathcal{R}_0}{\mathrm{d}s^*}\right|_{(1,1)} = \frac{\gamma\mathcal{P}}{\mu+\gamma} - 1.$$

Hence, a backward bifurcation occurs if $\mathcal{P} > \mathcal{P}_{\mathrm{crit}} = 1 + \mu/\gamma$. The value of $\mathcal{R}_0$ at the saddle node is the solution of

$$\mathcal{P}^2\mathcal{R}_0^2 + \left((2-\mathcal{P})\frac{\mu}{\mu+\gamma} - 1\right)2\mathcal{P}\mathcal{R}_0 + \left(1 - \mathcal{P}\frac{\mu}{\mu+\gamma}\right)^2 = 0.$$

A.4. A model with exogenous infection

Using Dulac's criterion,

$$\frac{\partial}{\partial s}\left(\frac{\mu - \beta si - \mu s}{si}\right) + \frac{\partial}{\partial i}\left(\frac{\mathcal{P}\beta ei + \nu e - \mu i}{si}\right) = -\frac{\mu}{s^2 i} - \frac{\mathcal{P}\beta}{s} - \frac{\nu(1-s)}{si^2} < 0.$$

Hence, no limit cycles exist and the non-trivial steady state is globally stable for $\mathcal{R}_0 > 1$. The steady-state value of i solves

$$\mathcal{P}\mathcal{R}_0^2 i^{*2} + \left(\frac{\nu}{\mu} + \mathcal{P}\frac{\nu}{\mu+\nu} - \mathcal{P}\mathcal{R}_0\right)\mathcal{R}_0 i^* + \frac{\nu^2(1-\mathcal{R}_0)}{\mu(\mu+\nu)} = 0.$$

Treating $\mathcal{R}_0$ as a function of i^*, differentiating and setting $(\mathcal{R}_0, i^*) = (1, 0)$, we obtain

$$\left.\frac{\mathrm{d}\mathcal{R}_0}{\mathrm{d}i^*}\right|_{(1,0)} = 1 + \frac{\mu}{\nu} - \frac{\mu^2\mathcal{P}}{\nu^2}.$$

Hence, a backward bifurcation occurs if $\mathcal{P} > \mathcal{P}_{\mathrm{crit}} = \nu(\mu+\nu)/\mu^2$.

A.5. A model with nonlinear transmission

Dulac's criterion does not apply here and there are some functions $h(i)$ that lead to limit cycles about the non-trivial steady state (Gomes *et al.* 2005, theorem 6). The non-trivial steady-state value of i solves

$$\frac{1}{\mathcal{R}_0(1 + h(i^*))} + \frac{\mu+\gamma}{\mu} i^* = 1.$$

Treating $\mathcal{R}_0$ as a function of i^*, differentiating and setting $(\mathcal{R}_0, i^*) = (1, 0)$, we obtain

$$\left.\frac{\mathrm{d}\mathcal{R}_0}{\mathrm{d}i^*}\right|_{(1,0)} = 1 + \frac{\gamma}{\mu} - h'(0).$$

Hence, a backward bifurcation occurs if $h'(0) > 1 + \gamma/\mu$.

A.6. Models with a carrier class

For these models (equations (3.4) or (4.1)), the proportion that become carriers upon recovery is an increasing function of the force of infection, λ. We define a variable $x = \lambda/\mu$ and set $q = q(x) \geq 0$, $q'(x) \geq 0$ for $x \geq 0$, and $\lim_{x\to\infty} q(x) = 1$. Hence, we have used μ to non-dimensionalize the argument of the function q. Steady states of equations (4.1) may be found as a function of x by substituting for s, setting $\mathrm{d}/\mathrm{d}t = 0$ and solving the resulting linear system. We obtain

$$c^* = \frac{\gamma q(x) i^*}{(1-k)\mu + \delta} \qquad r^* = \frac{\gamma}{\mu} i^* - (1-k)c^*,$$

and

$$\left(\frac{(\mu+\gamma)(1+x)}{\mu} + \frac{k\gamma q(x)x}{(1-k)\mu + \delta}\right) i^* = x.$$

This process determines a value for $\beta = \mu x/(i^* + \mathcal{P}c^*)$ and hence $\mathcal{R}_0$ may be expressed as a function of x allowing a bifurcation curve to be drawn. If there is no vertical transmission ($k = 0$), then

$$\mathcal{R}_0(x) = \frac{(\mu + \delta + \mathcal{P}\gamma q(0))(1+x)}{\mu + \delta + \mathcal{P}\gamma q(x)}.$$

A backward bifurcation at $(\mathcal{R}_0, i^*) = (1, 0)$ occurs if $\mathcal{P} > \mathcal{P}_{\mathrm{crit}} = (\mu+\delta)/(q'(0) - q(0))\gamma)$. Another special case is where $q(0) = 0$ and the eigenvalues of $\boldsymbol{K}$ are the matrix diagonals (see equation (4.2)). If $\beta/(\gamma+\mu) >$

$k\mu/(\delta+\mu)$, then

$$\mathcal{R}_0(x) = \frac{1 + x + \dfrac{\mu}{\gamma+\mu}\dfrac{k\gamma q(x)x}{(1-k)\mu+\delta}}{1+\mathcal{P}\dfrac{\gamma q(x)}{(1-k)\mu+\delta}}.$$

Note that $k\mu/(\delta+\mu)<1$, so above the threshold $\mathcal{R}_0=\beta/(\gamma+\mu)$.

Medley *et al.* (2001) used $q=f+(1-f)\exp(-0.645\lambda^{-0.455})$ for some constant f, $0\le f\le 1$. The two parameters given numerical values had previously been determined from data. As λ has units time^{-1} this is problematic, hence we chose $q(x)=\exp(-x^{-1})$ for the example presented in figure 2, being in the same spirit as the function used by Medley *et al.* (2001).

A.7. A model with two hosts

The model described by equations (4.3) is well posed in that $i_\mathrm{f}=0$ implies $(\mathrm{d}i_\mathrm{f}/\mathrm{d}t)\ge 0$ and $i_\mathrm{f}=1$ implies $(\mathrm{d}i_\mathrm{f}/\mathrm{d}t)<0$, with similar results for i_m. Limit cycles are not possible because

$$\frac{\partial}{\partial i_\mathrm{f}}(\beta_\mathrm{m}s_\mathrm{f}i_\mathrm{m}-\gamma_\mathrm{f}i_\mathrm{f})+\frac{\partial}{\partial i_\mathrm{m}}(\beta_\mathrm{f}s_\mathrm{m}i_\mathrm{f}-\gamma_\mathrm{m}i_\mathrm{m})$$
$$=-\beta_\mathrm{m}i_\mathrm{m}-\gamma_\mathrm{f}-\beta_\mathrm{f}i_\mathrm{f}-\gamma_\mathrm{m}<0.$$

Steady-state solutions solve

$$\begin{pmatrix}-\gamma_\mathrm{f} & \beta_\mathrm{m}s_\mathrm{f}\\ \beta_\mathrm{f}s_\mathrm{m} & -\gamma_\mathrm{m}\end{pmatrix}\begin{pmatrix}i_\mathrm{f}\\ i_\mathrm{m}\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

A non-trivial solution $(i_\mathrm{f},i_\mathrm{m})=(i_\mathrm{f}^*,i_\mathrm{m}^*)$ is possible only if the matrix is singular, hence $\beta_\mathrm{f}\beta_\mathrm{m}s_\mathrm{f}^*s_\mathrm{m}^*=\gamma_\mathrm{f}\gamma_\mathrm{m}$ and

$$i_\mathrm{f}^*=\frac{(\mathcal{R}_0)^2-1}{(\mathcal{R}_0)^2+\frac{\beta_\mathrm{f}}{\gamma_\mathrm{m}}}\qquad i_\mathrm{m}^*=\frac{(\mathcal{R}_0)^2-1}{(\mathcal{R}_0)^2+\frac{\beta_\mathrm{m}}{\gamma_\mathrm{f}}}$$

$$\mathcal{R}_0=\sqrt{\frac{\beta_\mathrm{f}\beta_\mathrm{m}}{\gamma_\mathrm{f}\gamma_\mathrm{m}}}.$$

The Jacobian matrix for the system is

$$\boldsymbol{J}(i_\mathrm{f},i_\mathrm{m})=\begin{pmatrix}-\gamma_\mathrm{f}-\beta_\mathrm{m}i_\mathrm{m} & \beta_\mathrm{m}s_\mathrm{f}\\ \beta_\mathrm{f}s_\mathrm{m} & -\gamma_\mathrm{m}-\beta_\mathrm{f}i_\mathrm{f}\end{pmatrix}.$$

The trace of $\boldsymbol{J}$ is always negative, hence a steady state is stable if the determinant is positive. At the infection-free steady state, $|\boldsymbol{J}(0,0)|=\gamma_\mathrm{f}\gamma_\mathrm{m}-\beta_\mathrm{m}\beta_\mathrm{f}$ and we have stability for $\mathcal{R}_0<1$. At the non-trivial steady state, $|\boldsymbol{J}(i_\mathrm{f}^*,i_\mathrm{m}^*)|>0$ and we have stability whenever the steady state exists, which is when $\mathcal{R}_0>1$.

A.8. A model with two hosts and two pathogens

Steady states of the model described by equations (4.4) solve

$$\begin{pmatrix}-\gamma_{\mathrm{f}1} & \beta_{\mathrm{m}1}s_\mathrm{f} & 0 & 0\\ \beta_{\mathrm{f}1}s_\mathrm{m} & -\gamma_{\mathrm{m}1} & 0 & 0\\ 0 & 0 & -\gamma_{\mathrm{f}2} & \beta_{\mathrm{m}2}s_\mathrm{f}\\ 0 & 0 & \beta_{\mathrm{f}2}s_\mathrm{m} & -\gamma_{\mathrm{m}2}\end{pmatrix}\begin{pmatrix}i_{\mathrm{f}1}\\ i_{\mathrm{m}1}\\ i_{\mathrm{f}2}\\ i_{\mathrm{m}2}\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

For coexistence of the two pathogens, it is necessary that $\beta_{\mathrm{f}1}\beta_{\mathrm{m}1}s_\mathrm{f}^*s_\mathrm{m}^*=\gamma_{\mathrm{f}1}\gamma_{\mathrm{m}1}$ and $\beta_{\mathrm{f}2}\beta_{\mathrm{m}2}s_\mathrm{f}^*s_\mathrm{m}^*=\gamma_{\mathrm{f}2}\gamma_{\mathrm{m}2}$, hence $\mathcal{R}_0^1=\mathcal{R}_0^2$. The steady-state solution with only one pathogen present is as in the previous section.

The Jacobian matrix for the system is

$$\begin{pmatrix}-\gamma_{\mathrm{f}1}-\beta_{\mathrm{m}1}i_{\mathrm{m}1} & \beta_{\mathrm{m}1}s_\mathrm{f} & -\beta_{\mathrm{m}1}i_{\mathrm{m}1} & 0\\ \beta_{\mathrm{f}1}s_\mathrm{m} & -\gamma_{\mathrm{m}1}-\beta_{\mathrm{f}1}i_{\mathrm{f}1} & 0 & -\beta_{\mathrm{f}1}i_{\mathrm{f}1}\\ -\beta_{\mathrm{m}2}i_{\mathrm{m}2} & 0 & -\gamma_{\mathrm{f}2}-\beta_{\mathrm{m}2}i_{\mathrm{m}2} & \beta_{\mathrm{m}2}s_\mathrm{f}\\ 0 & -\beta_{\mathrm{f}2}i_{\mathrm{f}2} & \beta_{\mathrm{f}2}s_\mathrm{m} & -\gamma_{\mathrm{m}2}-\beta_{\mathrm{f}2}i_{\mathrm{f}2}\end{pmatrix},$$

which, evaluated at the steady state $(i_{\mathrm{f}1},i_{\mathrm{m}1},i_{\mathrm{f}2},i_{\mathrm{m}2})=(i_{\mathrm{f}1}^*,i_{\mathrm{m}1}^*,0,0)$, is

$$\begin{pmatrix}-\gamma_{\mathrm{f}1}-\beta_{\mathrm{m}1}\mathrm{i}_{\mathrm{m}1}^* & \beta_{\mathrm{m}1}s_\mathrm{f}^* & -\beta_{\mathrm{m}1}i_{\mathrm{m}1}^* & 0\\ \beta_{\mathrm{f}1}s_\mathrm{m}^* & -\gamma_{\mathrm{m}1}-\beta_{\mathrm{f}1}i_{\mathrm{f}1}^* & 0 & -\beta_{\mathrm{f}1}i_{\mathrm{f}1}^*\\ 0 & 0 & -\gamma_{\mathrm{f}2} & \beta_{\mathrm{m}2}s_\mathrm{f}^*\\ 0 & 0 & \beta_{\mathrm{f}2}s_\mathrm{m}^* & -\gamma_{\mathrm{m}2}\end{pmatrix},$$

with $s_\mathrm{f}^*=1-i_{\mathrm{f}1}^*$ and $s_\mathrm{m}^*=1-i_{\mathrm{m}1}^*$. The four eigenvalues of this matrix are the two eigenvalues of each of the leading and trailing 2×2 submatrices on the diagonal. The leading matrix is stable if the steady state exists, the trailing matrix has a negative trace and the determinant is equal to

$$\gamma_{\mathrm{f}2}\gamma_{\mathrm{m}2}-\beta_{\mathrm{m}2}\beta_{\mathrm{f}2}s_\mathrm{f}^*s_\mathrm{m}^*=\gamma_{\mathrm{f}2}\gamma_{\mathrm{m}2}\left(1-\frac{\mathcal{R}_0^2}{\mathcal{R}_0^1}\right).$$

Hence, the steady state is stable whenever $\mathcal{R}_0^1>\mathcal{R}_0^2$.

A.9. A model with variable population density

If we represent the number of hosts in an area by N and the number of infected hosts by I, then the equations for this model are

$$\frac{\mathrm{d}N}{\mathrm{d}t}=(\eta(N)-\mu(N))N-\alpha I,$$
$$\frac{\mathrm{d}I}{\mathrm{d}t}=\beta(N)\frac{SI}{N}-(\alpha+\mu(N))I,$$

with $S+I=N$. The equation for the proportion of the population infected, $i=I/N$, is then $(5.1)_2$

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\frac{1}{N}\left(\frac{\mathrm{d}I}{\mathrm{d}t}-i\frac{\mathrm{d}N}{\mathrm{d}t}\right)=\beta(N)si-(\alpha s+\eta(N)).$$

The model described by equations (5.1) is a special case of that analysed by Roberts & Jowett (1996), who found four possible steady states: the trivial steady state $(N,i)=(0,0)$ and the infection-free steady state $(N,I)=(K,0)$, which exist for all parameter values; as well as host extinction and endemic steady states. The host extinction steady state has $N=0$ and $i=1-\eta(0)/(\beta(0)-\alpha)$, hence for $0<i<1$, it is necessary that $\beta(0)/(\alpha+\eta(0))>1$. At the endemic steady state, $\mathcal{R}(N^*)s^*=1$, where $\alpha i^*=\eta(N^*)-\mu(N^*)$ and $s^*+i^*=1$. A unique endemic steady-state solution exists if $\mathcal{R}_0>1$.

Using Dulac's criterion,

$$\frac{\partial}{\partial N}\left(\frac{(\eta(N)-\mu(N)-\alpha i)N}{Nsi}\right)+\frac{\partial}{\partial i}\left(\frac{\beta(N)si-(\alpha s+\eta(N))i}{Nsi}\right)$$

$$=\frac{\eta'(N)-\mu'(N)}{is}-\frac{\eta(N)}{Ns^2}<0,$$

and no limit cycles exist.

The Jacobian matrix for the system is

$$\begin{pmatrix} \eta(N)-\mu(N)-\alpha i \\ +(\eta'(N)-\mu'(N))N & -\alpha N \\ (\beta'(N)s-\eta'(N))i & (\beta(N)-\alpha)(s-i)-\eta(N) \end{pmatrix}.$$

Hence, the steady state $(N, i)=(0, 0)$ is unstable as long as $\eta(0)>\mu(0)$, which is essential for the population to grow away from zero. The steady state $(N, i)=(K, 0)$ is stable if $\mathcal{R}_0<1$ and unstable otherwise. The Jacobian at the host extinction steady state has eigenvalues $\eta(0)-\mu(0)-\alpha i$ and $-(\beta(0)-\alpha)i$. If this steady state exists, then $\beta(0)>\alpha$; so for it to be stable, it is necessary that $i>(\eta(0)-\mu(0))/\alpha$. The Jacobian matrix at the endemic steady state simplifies to

$$\begin{pmatrix} (\eta'(N^*)-\mu'(N^*))N^* & -\alpha N^* \\ (\beta'(N^*)s^*-\eta'(N^*))i^* & -(\beta(N^*)-\alpha)i^* \end{pmatrix},$$

which has negative trace and positive determinant. Hence, the endemic steady state is always stable when it exists.

A.10. A model with variable population density and two pathogens

If we write $\boldsymbol{x}=(N, i_1, i_2)'$, then equations (5.2) become

$$\frac{\mathrm{d}x_j}{\mathrm{d}t}=F_j(\boldsymbol{x})x_j,$$

for $j=1, 2, 3$, where

$$\boldsymbol{F}(\boldsymbol{x})=\begin{pmatrix} \eta(N)-\mu(N)-\alpha_1 i_1-\alpha_2 i_2 \\ \beta_1(N)(1-i_1-i_2)-\alpha_1-\eta(N)+\alpha_1 i_1+\alpha_2 i_2 \\ \beta_2(N)(1-i_1-i_2)-\alpha_2-\eta(N)+\alpha_1 i_1+\alpha_2 i_2 \end{pmatrix}.$$

This system has the potential for eight different steady states. These include four with $N=0$: the trivial steady state and host extinction states with either one or both pathogens present. Setting these aside, the steady states of interest are (i) $\boldsymbol{x}_0=(K, 0, 0)'$, the infection-free steady state where $F_1(\boldsymbol{x}_0)=\boldsymbol{x}_2=\boldsymbol{x}_3=0$, (ii) $\boldsymbol{x}_1=(N_1^*, i_1^*, 0)'$, the steady state with pathogen 1 only present where $F_1(\boldsymbol{x}_1)=F_2(\boldsymbol{x}_1)=\boldsymbol{x}_3=0$, (iii) $\boldsymbol{x}_2=(N_2^*, 0, i_2^*)'$, the steady state with pathogen 2 only present where $F_1(\boldsymbol{x}_2)=\boldsymbol{x}_2=F_3(\boldsymbol{x}_2)=0$, and (iv) $\boldsymbol{x}_3=(N^{**}, i_1^{**}, i_2^{**})'$, the coexistence steady state where $F_1(\boldsymbol{x}_3)=F_2(\boldsymbol{x}_3)=F_3(\boldsymbol{x}_3)=0$. The first three of these steady states may be found in the same way that those for equations (5.1) were found in the previous section. To find the coexistence steady state, we define $\mathcal{R}_j(N)=\beta_j(N)/(\alpha_j+\mu(N))$ for $j=1, 2$. For the steady state to exist, it is then necessary that $\mathcal{R}_1(N^{**})s^{**}=\mathcal{R}_2(N^{**})s^{**}=1$, which defines s^{**} and N^{**}. The steady-state infection prevalences are

$$i_1^{**}=\frac{\eta(N^{**})-\mu(N^{**})-\alpha_2(1-s^{**})}{\alpha_1-\alpha_2}$$

$$i_2^{**}=-\frac{\eta(N^{**})-\mu(N^{**})-\alpha_1(1-s^{**})}{\alpha_1-\alpha_2}.$$

The Jacobian matrix of the system is

$$\boldsymbol{J}(\boldsymbol{x})=\begin{pmatrix} x_1\frac{\mathrm{d}F_1}{\mathrm{d}x_1}+F_1(\boldsymbol{x}) & x_1\frac{\mathrm{d}F_1}{\mathrm{d}x_2} & x_1\frac{\mathrm{d}F_1}{\mathrm{d}x_3} \\ x_2\frac{\mathrm{d}F_2}{\mathrm{d}x_1} & x_2\frac{\mathrm{d}F_2}{\mathrm{d}x_2}+F_2(\boldsymbol{x}) & x_2\frac{\mathrm{d}F_2}{\mathrm{d}x_3} \\ x_3\frac{\mathrm{d}F_3}{\mathrm{d}x_1} & x_3\frac{dF_3}{dx_2} & x_3\frac{\mathrm{d}F_3}{\mathrm{d}x_3}+F_3(\boldsymbol{x}) \end{pmatrix}.$$

At the infection-free steady state, $\boldsymbol{x}_0=(K, 0, 0)'$,

$$\boldsymbol{J}(x_0)=\begin{pmatrix} (\eta'(K)-\mu'(K))K & -\alpha_1 K & -\alpha_2 K \\ 0 & \beta_1(K)-\alpha_1-\mu(K) & 0 \\ 0 & 0 & \beta_2(K)-\alpha_2-\mu(K) \end{pmatrix}.$$

Hence, this steady state is locally stable if $\mathcal{R}_1(K)<1$ and $\mathcal{R}_2(K)<1$. At the steady state with only pathogen variant 1 present, $\boldsymbol{x}_1=(N_1^*, i_1^*, 0)'$

$$\boldsymbol{J}(\boldsymbol{x}_1)$$

$$=\begin{pmatrix} (\eta'(N_1^*)-\mu'(N_1^*))N_1^* & -\alpha_1 N_1^* & -\alpha_2 N_1^* \\ (\beta_1'(N_1^*)s_1^*-\eta'(N_1^*))i_1^* & (\alpha_1-\beta_1(N_1^*))i_1^* & (\alpha_2-\beta_1(N_1^*))i_1^* \\ 0 & 0 & \beta_2(N_1^*)-\alpha_2-\mu(N_1^*) \end{pmatrix}.$$

As $\mathcal{R}_1(N_1^*)s_1^*=1$, $\beta_1(N_1^*)>\alpha_1$ and the leading 2×2 submatrix of $\boldsymbol{J}(\boldsymbol{x}_1)$ have negative eigenvalues. Hence, this steady state exists if $\mathcal{R}_1(K)>1$ and is locally stable if $\mathcal{R}_2(N_1^*)<1$. Similarly, the steady state with strain 2 only present exists if $\mathcal{R}_2(K)>1$ and is locally stable if $\mathcal{R}_1(N_2^*)<1$. Numerical results suggest that the coexistence steady state $\boldsymbol{x}_3=(N^{**}, i_1^{**}, i_2^{**})'$ is stable whenever it exists. Andreasen & Pugliese (1995) proved local stability of the coexistence equilibrium of the special case of this model that they analysed.

REFERENCES

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control*. Oxford, UK: Oxford University Press.

Andreasen, V. & Pugliese, A. 1995 Pathogen coexistence induced by density-dependent host mortality. *J. Theor. Biol.* **177**, 159–165.

Ball, F. & Becker, N. G. 2006 Control of transmission with two types of infection. *Math. Biosci.* **200**, 170–187. (doi:10.1016/j.mbs.2005.12.024)

Blower, S. M., McLean, A. R., Porco, T. C., Small, P. M., Hopewell, P. C., Sanchez, M. A. & Moss, A. R. 1995 The intrinsic transmission dynamics of tuberculosis epidemics. *Nat. Med.* **1**, 815–821. (doi:10.1038/nm0895-815)

Dieckmann, U. & Metz, J. A. J. 2006 Surprising evolutionary predictions from enhanced ecological realism. *Theor. Popul. Biol.* **69**, 263–281. (doi:10.1016/j.tpb.2005.12.001)

Dieckmann, U., Metz, J. A. J., Sabelis, M. W. & Sigmund, K. 2002 *Adaptive dynamics of infectious diseases*. Cambridge, UK: Cambridge University Press.

Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. Wiley series in mathematical and computational biology*. Chichester, UK: Wiley.

Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382. (doi:10.1007/BF00178324)

Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1995 The legacy of Kermack and McKendrick. In *Epidemic models: their structure and relation to data* (ed. D. Mollison), pp. 95–115. Cambridge, UK: Cambridge University Press.

Edelstein-Keshet, L. 2005 *Mathematical models in biology*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Esteva, L. & Vargas, C. 1998 Analysis of a dengue disease transmission model. *Math. Biosci.* **150**, 131–151. (doi:10.1016/S0025-5564(98)10003-2)

Esteva, L. & Vargas, C. 2000 Influence of vertical and mechanical transmission on the dynamics of dengue disease. *Math. Biosci.* **167**, 51–64. (doi:10.1016/S0025-5564(00)00024-9)

Feng, Z. & Velasco-Hernandez, J. X. 1997 Competitive exclusion in a vector–host model for the dengue fever. *J. Math. Biol.* **35**, 523–544. (doi:10.1007/s002850050064)

Feng, Z., Castillo-Chavez, C. & Capurro, A. F. 2000 A model for tuberculosis with exogenous reinfection. *Theor. Popul. Biol.* **57**, 235–247. (doi:10.1006/tpbi.2000.1451)

Gomes, M. G. M., Franco, A. O., Gomes, M. C. & Medley, G. F. 2004 The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proc. R. Soc. B* **271**, 617–623. (doi:10.1098/rspb.2003.2606)

Gomes, M. G. M., Margheri, A., Medley, G. F. & Rebelo, C. 2005 Dynamical behaviour of epidemiological models with sub-optimal immunity and nonlinear incidence. *J. Math. Biol.* **51**, 414–430. (doi:10.1007/s00285-005-0331-9)

Grassly, N. C. & Fraser, C. 2006 Seasonal infectious disease epidemiology. *Proc. R. Soc. B* **273**, 2541–2550. (doi:10.1098/rspb.2006.3604)

Haydon, D., Chase-Topping, M., Shaw, D. J., Matthews, L., Friar, J., Wilesmith, J. & Woolhouse, M. E. J. 2003 The construction and analysis of epidemic trees with reference to the 2001 UK foot-and-mouth outbreak. *Proc. R. Soc. B* **270**, 121–127. (doi:10.1098/rspb.2002.2191)

Heesterbeek, J. A. P. 2002 A brief history of R_0 and a recipe for its calculation. *Acta Biotheor.* **50**, 189–204. (doi:10.1023/A:1016599411804)

Heesterbeek, J. A. P. & Roberts, M. G. 1995*a* Threshold quantities for helminth infections. *J. Math. Biol.* **33**, 415–434. (doi:10.1007/BF00176380)

Heesterbeek, J. A. P. & Roberts, M. G. 1995*b* Threshold quantities for infectious diseases in periodic environments. *J. Biol. Syst.* **3**, 779–787. (doi:10.1142/S021833909500071X)

Heesterbeek, J. A. P. & Roberts, M. G. 2007 The type-reproduction number T in models for infectious disease control. *Math. Biosci.* **206**, 3–10. (doi:10.1016/j.mbs.2004.10.013)

Heffernan, J. M., Smith, R. J. & Wahl, L. M. 2005 Perspectives on the basic reproductive ratio. *J. R. Soc. Interface* **2**, 281–293. (doi:10.1098/rsif.2005.0042)

Jordan, D. W. & Smith, P. 1999 *Nonlinear ordinary differential equations*, 3rd edn. Oxford, UK: Oxford University Press.

Kao, R. R. 2006 Evolution of pathogens towards low R_0 in heteregeneous populations. *J. Theor. Biol.* **242**, 634–642. (doi:10.1016/j.jtbi.2006.04.003)

Kermack, W. O. & McKendrick, A. G. 1927 A Contribution to the mathematical theory of epidemics. I. *Proc. R. Soc. A* **115**, 700–721. (Reprinted with parts II. and III. in *Bulletin of Mathematical Biology*, **53** (1991) 33–118.) (doi:10.1098/rspa.1927.0118)

Kiss, I. Z., Green, D. M. & Kao, R. R. 2006 Infectious disease control using contact tracing in random and scale-free networks. *J. R. Soc. Interface* **3**, 55–62. (doi:10.1098/rsif.2005.0079)

May, R. M. 2006 Network structure and the biology of populations. *Trends Ecol. Evol.* **21**, 394–399. (doi:10.1016/j.tree.2006.03.013)

Medley, G. F., Lindop, N. A., Edmunds, W. J. & Nokes, D. J. 2001 Hepatitis-B virus endemicity: heterogeneity, catastrophic dynamics and control. *Nat. Med.* **5**, 619–624. (doi:10.1038/87953)

Roberts, M. G. 1996 Bovine tuberculosis in possum populations, and its eradication or control by culling or vaccination. *J. Anim. Ecol.* **65**, 451–464. (doi:10.2307/5780)

Roberts, M. G. & Dobson, A. P. 1995 The population dynamics of communities of parasitic helminths. *Math. Biosci.* **126**, 191–214. (doi:10.1016/0025-5564(94)00036-Y)

Roberts, M. G. & Heesterbeek, J. A. P. 2003 A new method for estimating the effort required to control infectious diseases. *Proc. R. Soc. B* **270**, 1359–1364. (doi:10.1098/rspb.2003.2339)

Roberts, M. G. & Jowett, J. 1996 An SEI model with density dependent demographics and epidemiology. *IMA J. Math. Appl. Med. Biol.* **13**, 245–257.

Roberts, M. G. & Kao, R. R. 1998 The dynamics of an infectious disease in a population with birth pulses. *Math. Biosci.* **149**, 23–36. (doi:10.1016/S0025-5564(97)10016-5)

Roberts, M. G. & Tobias, M. I. 2000 Predicting and preventing measles epidemics in New Zealand: application of a mathematical model. *Epidemiol. Infect.* **124**, 279–287. (doi:10.1017/S0950268899003556)

Roberts, M. G., Baker, M., Jennings, L. C., Sertsou, G. & Wilson, N. 2007 A model for the spread and control of pandemic influenza in an isolated geographical region. *J. R. Soc. Interface* **4**, 325–330. (doi:10.1098/rsif.2006.0176)

Safan, M., Heesterbeek, J. A. P. & Dietz, K. 2006 The minimum effort required to eradicate infections in models with backward bifurcation. *J. Math. Biol.* **53**, 703–718. (doi:10.1007/s00285-006-0028-8)

Soewono, E. & Supriatna, A. K. 2001 A two-dimensional model for the transmission of dengue fever disease. *Bull. Malaysian Math. Sci. Soc.* **24**, 49–57.

Strogatz, S. H. 1994 *Nonlinear dynamics and chaos*. Cambridge, MA: Westview Press.

van den Driessche, P. & Watmough, J. 2002 Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29–48. (doi:10.1016/S0025-5564(02)00108-6)